

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 10:39:49 ; Search time 186.16 Seconds  
(without alignments)  
9305.869 Million cell updates/sec

Title: US-09-784-340-1

Perfect score: 2759  
Sequence: 1 caaccatgcagatcagctgtc.....ctgtcagccgttaagctagc 2759

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

N\_Geneseq\_0601:\*

1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:\*

2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:\*

3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:\*

4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:\*

5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:\*

6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:\*

7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:\*

8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:\*

9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:\*

10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT:\*

11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:\*

12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:\*

13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:\*

14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:\*

15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:\*

16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:\*

17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:\*

18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:\*

19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:\*

20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:\*

21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	967.4	35.1	1650	21	AAC65396	Human carboxylate
2	780.4	28.3	1976	21	AAZ95206	Human UDP-glucuron
3	775.6	28.1	2107	19	AAV15900	Uridine diphospho-
4	766	27.8	2092	21	AAZ95199	Human UDP-glucuron
5	750	27.2	1854	21	AAZ95200	Human UDP-glucuron
6	486.4	17.6	515	20	AAV87412	EST clone BR77. H
7	350	12.7	350	21	AAC03286	Human secreted pro
8	317.8	11.5	2368	13	AAQ27370	HUG-Br2. Homo sap
9	316.6	11.5	2351	13	AAQ27369	HUG-Br1. Homo sap
10	272.2	9.9	735	21	AAZ45118	UDP-glucuronosyltr
11	200.2	7.3	2312	21	AAZ95207	Human UGT2B15 exon

12	196.4	7.1	936	22	AAZ58252 Oligonucleotide D1
13	196.4	7.1	936	22	AAZ58254 Oligonucleotide D1
14	196.4	7.1	936	22	AAZ58257 Oligonucleotide D1
15	196.4	7.1	936	22	AAZ58259 Oligonucleotide D2
16	196.4	7.1	936	22	AAZ58262 Oligonucleotide D2
17	196.4	7.1	938	22	AAZ58255 Oligonucleotide D1
18	195.8	7.1	1323	21	AAZ95193 Human UGT2B4 exon
19	195.4	7.1	556	21	AAZ95205 Human UGT2B7 exon
20	194.2	7.0	936	22	AAZ58252 Oligonucleotide D1
21	194.2	7.0	936	22	AAZ58254 Oligonucleotide D1
22	194.2	7.0	936	22	AAZ58257 Oligonucleotide D1
23	194.2	7.0	936	22	AAZ58259 Oligonucleotide D2
24	194.2	7.0	936	22	AAZ58262 Oligonucleotide D2
25	194.2	7.0	938	22	AAZ58255 Oligonucleotide D1
26	193.8	7.0	938	22	AAZ58257 Oligonucleotide D1
27	188.4	6.5	1589	21	AAZ95211 Human UGT2B4 exon
28	179.8	6.5	1686	21	AAZ95201 Human UGT2B7 exon
29	145.6	5.3	391	20	AAZ95208 Human lung tumour
30	145.6	5.3	391	21	AAZ95206 Human UGT2B15 exon
31	144.8	5.2	1602	21	AAZ95210 Human UGT2B7 exon
32	144.6	5.2	1591	21	AAZ95204 Human UGT2B7 exon
33	143.2	5.2	2320	21	AAZ58107 Human PRO1780 (UNG)
34	143.2	5.2	2320	21	AAZ37101 Human PRO1780 (UNG)
35	143.2	5.2	2320	22	AAZ54401 Primer #82 used in
36	141.6	5.1	2341	22	AAZ93775 Human cDNA encodin
37	138.6	5.0	689	21	AAZ95197 Human UGT2B4 exon
38	132.8	4.8	283	21	AAZ87467 Rat hepatocyte car
39	131	4.7	223	21	AAZ87255 Rat hepatocyte car
40	120.2	4.4	1020	21	AAZ95208 Human UGT2B15 exon
41	114	4.1	164	21	AAZ87154 Rat hepatocyte car
42	110.2	4.0	273	21	AAZ87491 Rat hepatocyte car
43	109.8	4.0	746	21	AAZ95194 Human UGT2B4 exon
44	101.4	3.7	1340	21	AAZ95202 Human UGT2B7 exon
45	98.6	3.6	1822	21	AAZ95203 Human UGT2B7 exon

### ALIGNMENTS

RESULT	1
AAC65396	AAC65396 standard; cDNA; 1650 BP.
ID	XX
AC	AAC65396;
DT	13-FEB-2001 (first entry)
XX	XX
DE	Human carboxylate-modifying enzyme cDNA Incyte ID No: 2912330CB1.
XX	XX
KW	Human; carboxylate-modifying enzyme; CME; antidiabetic;
KW	Immunosuppressive; anti-HIV; antiinflammatory; antianaemic;
KW	antisthmatic; antiarteriosclerotic; antihypoid; hepatotropic;
KW	nephrotropic; antiout; thymimetic; neuroprotective; osteopathic;
KW	arthritic; antipruritic; uropathic; ophthalmological;
KW	dermatological; antitumor; cytostatic; vituicide; antibacterial;
KW	fungicide; protozoicide; tranquiliser; vulnary; diabetes;
KW	autoimmune disorder; inflammatory disorder; infection; ss.
XX	XX
OS	Homo sapiens.
XX	XX
PN	W0200063351-A2.
XX	XX
PD	26-OCT-2000.
XX	XX
PF	20-APR-2000; 2000MO-US10882.
XX	XX
PR	21-APR-1999; 99US-0130383.
XX	XX
PA	(INCY-) INCYTE GENOMICS INC.
PI	Lal P, Yue H, Tang YT, Hillman JL, Baughn MR, Yang J;
XX	XX
DR	WPI; 2000-672729/65.

Done UDP glucuronosyl transferases

DR P-PSDB: AAB28677.  
 XX Novel carbohydrate modifying enzyme polypeptides and polynucleotides  
 PT for diagnosis, treatment and prevention of carbohydrate metabolism  
 PT disorders, autoimmune/inflammatory disorders, and cancer  
 XX  
 PS Claim 4; Page 75; 75pp; English.  
 XX  
 CC The present cDNA sequence encodes a human carbohydrate-modifying enzyme  
 CC (CME). CME polynucleotides and polypeptides are useful for treating and  
 CC diagnosing diseases associated with CME such as diabetes,  
 CC autoimmune/inflammatory disorders such as AIDS, Addison's disease,  
 CC adult respiratory distress syndrome, allergies, anaemia, asthma,  
 CC atherosclerosis, autoimmune thyroiditis, bronchitis, cholecystitis,  
 CC contact dermatitis, Crohn's disease, emphysema, erythroblastosis fetalis,  
 CC glomerulonephritis, Good pasture's syndrome, gout, Grave's disease,  
 CC Hashimoto's thyroiditis, multiple sclerosis, myasthenia gravis,  
 CC osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis,  
 CC Reiter's syndrome, arthritis, scleroderma, Sjogren's syndrome, systemic  
 CC lupus erythematosus, ulcerative colitis, uveitis, Werner syndrome,  
 CC complications of cancer, haemodialysis, and extracorporeal circulation,  
 CC viral, bacterial, fungal parasitic, protozoal, and helminthic infections,  
 CC trauma, or cancer. CME, or its catalytic or immunogenic fragment, is  
 CC useful for drug screening.  
 CC  
 XX  
 SO Sequence 1650 BP; 489 A; 330 C; 354 G; 477 T; 0 other;

Query Match 35.1%; Score 967.4; DB 21; Length 1650;  
 Best Local Similarity 75.7%; Pred. No. 3, 3e-187;  
 Matches 1228; Conservative 0; Mismatches 386; Indels 9; Gaps 2;

QY 34 catcatgaagtcgacagtcagcttgatcttgctcctgcagctctctgt---gt 90  
 DB 28 catcagagatgctatgaatgaatgactcagctctctcctcatagcagtgctgactt 87  
 QY 91 tggctgtgactcttgagggaagtcctggtgtgctgctgacatgagcattgcttaa 150  
 DB 88 tactcttggaagtcgtggaagtcgtggtgtggtccacgaatcgaacacgtgagtaa 147  
 QY 151 tgcacagtcattctgaagagctcagtagtgagagccatgagtaacagtaactgacta 210  
 DB 148 tataagacactcctgataacttgctccagagagtcagtgagtgactgtgactc 207  
 QY 211 ctcaagcctctgtaattgactacaggaagcctctgctcaattgaattgagtgctca 270  
 DB 208 ttcagcttcattcttctgactccacagccatctactcttaattgaattatcc 267  
 QY 271 tatgcacagagacagacaggaagaataattgtgactgactgcttga-----a 324  
 DB 268 tgcattcttaactaaacgtgattgagatataccaagcagctggttaagagatgggc 327  
 QY 325 tgccttgcaagcttcaactctgacatcagttataaataatgaatttttggttga 384  
 DB 328 gaaactccaagaagacacttggctacatattttcacaagaatcaagaatacatgagac 387  
 QY 385 aataagaggaactttaaataatgatgtgtgagagcttactactcaacatacagctta 444  
 DB 388 atttaatgacatactatagaagatcgtgaagataatagtttcaaaataagaactta 447  
 QY 445 gaagctacaggaacccaactacagatgtaattatagaccctgtagtccctgtgaga 504  
 DB 448 gaaactacagagtgcaagatttgattgttcttgcaagtgctgcttccctttgtgta 507  
 QY 505 cctgctggtgctgctgctgctgctcctttgtgctcaacttagaattctgtlaggag 564  
 DB 508 gctgctggtgctgctgctgctgctcctttgtgctcaacttagaattctgtlaggag 567  
 QY 565 caataatgagcgaagctgtgaggaactcagctcagctcagctcagctcagctcagct 624  
 DB 568 cgaattgaaagcagtagtgagagcttctgctcctcctctctatgctgctgctgctat 627  
 QY 625 gacagagctaacagacagatgaccttctggaagagtaaaataatcaatgcttccagt 684

DB 628 gtccgaactagtcgacaaatgacttcaatagagaggttaaaataatgactatgct 687  
 QY 685 ttgttccactctgagatcagatgaactaactatcttggggaaggttatagtaa 744  
 DB 688 ttatttgaatttggctcaaatatttgataagaagagtgagatcagcttccatgta 747  
 QY 745 ggcattggaagcccaactacatattatgtgagactgtggaagaagctgagata 804  
 DB 748 agtcttagaagaccactacatgactatcgtagacatgagcaaaagctgacata 807  
 QY 805 acgaactatgggatttgaatttcttccaaacatatacactaacttgaattgttgg 864  
 DB 808 tcgaactactggtgatttcaacttccaccacacttccacaaatgttgaattgttgg 867  
 QY 865 aggatgactglaaacctgccaagcttgcctcaagaagataatgtgtccagag 924  
 DB 868 aggactcactgcaaacctgccaacccctaccgaaggaatggaagattgttccagag 927  
 QY 925 ttcaggggaagatgattgtgtgttcttctgtggtgctacttcaaaatgttaca 984  
 DB 928 ctctggaagaatgtgtgtgtgttcttctgtggtgctacttcaaaatgttaca 987  
 QY 985 agaaaagctaatatcatcttccagcctgcccagatcccaagaggtgtatagag 1044  
 DB 988 agaaagggccaatgtatttgcacagccttgcgaagatcccaaaagttctgtgag 1047  
 QY 1045 gtacaaaggaataaacacatccacataggaagccaactcgtgtatagttgatacc 1104  
 DB 1048 atttgatgaaataaacaagatactttagactcaactacactcgtgtataagtgatacc 1107  
 QY 1105 ccgaatgactctctgtgtatcccaaaccaagcttataactatcgtgttgaatgaa 1164  
 DB 1108 ccgaatgactctctgtgtatcccaaaccaagcttataactatcgtgttgaatgaa 1167  
 QY 1165 tgggactatgaagcattatccatgagtggtccctatgtgtgagttcccatattgtga 1224  
 DB 1168 tgggactatgaagcattatccatgagtggtccctatgtgtgagttcccatattgtga 1227  
 QY 1225 tcaagctgataacatgctcacaatgaagcgaagagcagcgtgtgaataaactcaa 1284  
 DB 1228 tcaagctgataacatgctcacaatgaagcgaagagcagcgtgtgaataaactcaa 1287  
 QY 1285 aactatgacaagcgaagatttactgagtggtcttgaagaacatgataccgattccctta 1344  
 DB 1288 aactatgacaagcgaagatttactgagtggtcttgaagaacatgataccgattccctta 1347  
 QY 1345 taaaggaatgctatgagattatcaagaattcaacatgacacactgtataagccctaga 1404  
 DB 1348 taaaggaatgctatgagattatcaagaattcaacatgacacactgtataagccctaga 1407  
 QY 1405 tcgaagcactcttctgagatgagtttgcacatgcccagaagagccaaagcacttgatc 1464  
 DB 1408 tcgaagcactcttctgagatgagtttgcacatgcccagaagagccaaagcacttgatc 1467  
 QY 1465 agctgccatgacccactcactggttccagacactactatagatgtgattggttccgtc 1524  
 DB 1468 agctgccatgacccactcactggttccagacactactatagatgtgattggttccgtc 1527  
 QY 1525 gacccgtgtggaacactgctatcttctgttcaaaaatgttttatttcttctgtc 1584  
 DB 1528 gacccgtgtggaacactgctatcttctgttcaaaaatgttttatttcttctgtc 1587  
 QY 1585 atttaataaactagaaagatagaagaaggggaatagatcttccaatccaagaagac 1644  
 DB 1588 atttaataaactagaaagatagaagaaggggaatagatcttccaatccaagaagac 1647  
 QY 1645 ctg 1647  
 DB 1648 ctg 1650

RESULT 2

AA295206  
ID AA295206 standard; DNA: 1976 BP.  
XX  
AC AA295206;  
XX  
DT 05-JUN-2000 (first entry)  
XX  
DE Human UDP-glucuronosyltransferase 2B15 nucleotide sequence.  
PS  
XX  
KW UDP-glucuronosyltransferase 2B15; UGT2B15; polymorphism; metabolism;  
KM drug interaction; detect; human; single nucleotide polymorphism;  
SNPs; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200006776-A1.  
XX  
PD 10-FEB-2000.  
XX  
PE 22-JUL-1999; 99MO-US16675.  
XX  
PR 28-JUL-1998; 98US-0094391.  
XX  
PA (AXYS-) AXYS PHARM INC.  
XX  
PI Galvin M, Miller A, Penny L, Riedy M;  
DR MPI; 2000-195321/17.  
DR P-PSDB; AAY78935.  
XX  
PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for  
PT genotyping individuals to predict rate of metabolism of substrates and  
PT for identifying potential drug interactions  
XX  
XX  
PS Disclosure; Page 56-59; 72pp; English.  
XX  
XX  
CC This sequence represents the human UDP-glucuronosyltransferase 2B15  
CC (UGT2B15) gene. UDP-glucuronosyltransferase (UGTs) are a family of  
CC enzymes that catalyse the glucuronic acid conjugation of a wide range of  
CC endogenous and exogenous substrates. The UGT2B gene subfamily encode  
CC steroid metabolizing isoforms in the liver. Alteration of the expression  
CC or function of UGTs may effect drug metabolism. The invention relates to  
CC non-chromosomal nucleic acid molecules, which comprise human UGT2B  
CC sequence polymorphisms (see AA295051-295110). Probes which detect the  
CC UGT2B locus polymorphisms can be used to detect altered UGT2B metabolism  
CC of a substrate in an individual. The nucleic acid molecules comprising a  
CC human UGT2B sequence polymorphism can be used in screening assays for  
CC genotyping individuals, also to predict their rate of metabolism of  
CC UGT2B substrate, potential drug-drug interactions and adverse side  
CC effects. The polymorphisms can be used as single nucleotide polymorphisms  
CC (SNPs) for detecting genetic linkage related to phenotypic variation in  
CC activity or expression of UGT2B protein. The polymorphism containing  
CC nucleic acid molecules may also be used for generating genetically  
CC modified non-human animals and for obtaining site specific gene  
CC modification in cell lines.  
XX  
XX  
SO Sequence 1976 BP; 594 A; 368 C; 419 G; 595 T; 0 other;

Query Match 28.3%; Score 780.4; DB 21; Length 1976;  
Best Local Similarity 68.5%; Pred. No. 2.7e-149;  
Matches 1113; Conservative 0; Mismatches 501; Indels 12; Gaps 2;

QY 35 atatgaagctgcacaaagcttgatctgtctcctgcagctctctgt--gtc 91  
DB 5 accaagaatcctctgaatgaagctcagctctctctctgatacagccagtgttaactt 64  
QY 92 ggcctgtgagatctctgtggaaagctcctgtgttgccctgtgacatgagccattgcttaatt 151  
DB 65 agctctggaagcgtgtgaaagctgtagtgtgtgcccacagaataacagcattgataaat 124  
QY 152 gtcaaggtcatctctagaagcctcatagtgaagagccatgaggttaacagtattgaactcac 211

DB 125 atgaagaacaatcctcgtgaagagcttgctcagagggtcatagaggtgactgtgtgacaatc 184  
QY 212 tcaagccctcgttaattgactacagaagcctctgtacattgaatttgaggtgtccat 271  
DB 185 tcgctcttactactctgtcaatgcagtaaatcatctgtctataataatgaaatttaccct 244  
QY 272 atgcc-----acagagcagaaacaggaagaaatgaaattgttgccttagctcgtg 322  
DB 245 acatcctttaaactaaatgtatttggaagatctctctcttgaaatctccgataagatgata 304  
QY 323 aatgtctgcagagctcatatcaacctgtgcaatcagttataaataatgattttttgtt 382  
DB 305 tatgtgtttcacaataaataacattttgttcataattttccaataatcacagattgtgttg 364  
QY 383 gaataaaggaaccttaaaatgatgtgtgagagctttatctacaatcacagcgttatg 442  
DB 365 gaattattgactacagtaacaagctctgtaaagatgcagtttgaataagaacaattatg 424  
QY 443 aagaagctacagaagaacaactacagatgatatgttatagacctgtgactcctgtgga 502  
DB 425 atgaactacaagaagltcaaatgttgatgtcatctctgtgcagatgcccataacctgttgc 484  
QY 503 gacctgagtgcgtgactgtgcagtcctcttctgtgtgcacacttaagaattctgttaga 562  
DB 485 gagctactgctgaactatttaacataacctctctgtgacagctctgactctctgttgc 544  
QY 563 ggcataatgagcgaagctgtgggaaactccagctccacttccctactgtactgtgct 622  
DB 545 tacacattgagaagaatgtgtgagattctctgttccctctcctctactgtacactgtgtt 604  
QY 623 atgacagagactaaagacagaatgaccttctctggaagagtaaaatcaatgacttca 682  
DB 605 atgcaagaatlaagtgatacaaatgatttcaatgagaagagataaaacatgatacaatg 664  
QY 683 gttgttccactctgtgatacagagatacagactatactattttggagaaggtttatagt 742  
DB 665 ctattatttgacctttgtgttccaattatgactcgtgaagaagtgagccagttttatagt 724  
QY 743 aagcattgaagaagcccactacaattatgtgagactgtgggaaagctgagataatgcta 802  
DB 725 gaagttctaagaagaccactacattatgtgagaatgggaaagctgaaatgtgctc 784  
QY 803 atacgaacatattggatttgatttccctcaacacataccaacttaactttgattgtt 862  
DB 785 attcgaacctatttggatttggatttccctgcgcacttccacttaacaattgttattgtt 844  
QY 863 ggaagattgactgttaacctgtccaaagcttgcctaagggaatgaaattgttcag 922  
DB 845 ggaagacttcaactgttaaaccaagccaaacctgtcctaaggaatggaaggtttgtgcag 904  
QY 923 agttcaggggaagatggtatgtgtgtttctctctgggtgcactgtttcaaatgttaca 982  
DB 905 agctctggaagaataatgttatgtgtgtttctctctgggtgcagatgacagtaacatgtca 964  
QY 983 gaagaaaggctataatcatctgtccagccctgtgccagatcccaagaaggtgtatgg 1042  
DB 965 gaagaaagtgccaacatgatgtcatcgccctgtgccagatcccaagaaggtgttatgg 1024  
QY 1043 aggtacaaagaaataaacatccatctagggagccatctcgtcgttatgttggata 1102  
DB 1025 agattgtatgcaagaagcccaaatcatttagttccaatccagactcgttcaagttgta 1084  
QY 1103 ccccaaatgatctctctgtgtcatcccaaaacaaagctttatcaactctgtgtgaaatg 1162  
DB 1085 ccccaagaatgaccttctgtgtcatcccaaaacaaagctttatcaactctgtgtgaaac 1144  
QY 1163 aatggatctataagactataacatgaggttccctatggttggaggttcccaatttgtt 1222  
DB 1145 aatgcatctatagagcagatccatcagatggtatccctatggttggcattccctgttgcg 1204  
QY 1223 gatcagcttgatatacagatgctacatgaagggcgaagggagcagctgtlagaataaacttc 1282  
DB 1205 gatcaacatgataaactgtctcacatgaaagccaaaggagcaccctcagctgttgacaatc 1264



QY	803	atcgcgaacatattggtgatttggaaatttcttccaacataccaactaacttggattt	862
Db	826	attcgacacatttggatttggaaatttcttccaacataccaactaacttggattt	885
QY	863	ggaggaatgcattgcaaacctgcgaagcttgccttaaggaaatggaaatttggcag	922
Db	886	ggaggaatgcattgcaaacctgcgaagcttgccttaaggaaatggaaatgtgtcag	945
QY	923	agttcaggggaaagatgatttggttgtttctctcgtgggttactgtttcaaatgttaca	982
Db	946	agctctgtagaaaatgtgatttggttgtttctctcgtgggttactgttcaaatgttaca	1005
QY	983	gaagaaaaggtcctaataattgtctcagccttgcagatccacagaagtgattatg	1042
Db	1006	gaagaaaaggtcctaataattgtctcagccttgcagatccacagaagtgattatg	1065
QY	1043	aggtcaagaaggaaaaacatccacatttaaggagcaatactcgtgtgattgata	1102
Db	1066	agatttgcagaaagaagcacaatacttlaagttccaatactgcattgataatgtgta	1125
QY	1103	ccccgaatgatctcttctgttcattcccaaaaccaaagcttatacactcagtgagatg	1162
Db	1126	ccccgaatgaccttcttgcgtatcccaaaaccaaagcttatacactcagtgagaaac	1185
QY	1163	aatgggattctatgaagctattaccatgagggtccctatggtggaggttcccaattggt	1222
Db	1186	aatgggattctatgagcgatctaccatggatccctatggtggaggttccctatggtcg	1245
QY	1223	gatcgacttgataaatatgcttcacatgtaagcccaagagagagcgtgtgtaataacttc	1282
Db	1246	gatcaacatgataaatatgcttcacatgtaagcccaagagagagcgtgtgtaataacttc	1305
QY	1283	aaaactatgacacagcgaaagatttacttggggcttggagaaacagatcattccgattcctct	1342
Db	1306	aggaccatgctcaagtagagatttgcctcaatgcattggaatcagatcattcaatgacccatc	1385
QY	1343	tataaagaagaatgctatgagattatccaagaattccaatgatacactgtataagccctta	1402
Db	1366	tataaagaagaatgctatgagattatccaagaattccaatgatacactgtataagcccttg	1425
QY	1403	gatcgagcagctctcttgcgtatcgagttgtcatgcccagaagagccagacccttgcga	1462
Db	1426	gatcgagcagctctcttgcgtatcgagttgtcatgcccagaagagccagacccttgcgg	1485
QY	1463	tcaagtcgccatgagcactacactgttccagacactactatagaattgattgggttcctg	1522
Db	1486	gtcgcgaagcccaaacctcaccctgggacccagatcacccctcttggatgtagatgacatccctg	1565
QY	1523	ctgacctgtgtgtgcacacgcgtatattctctgtgtccaaaatgtttttatttcttcgtcaa	1582
Db	1546	ctgacctgtgtgtgcacacatgatatattatgatacaaaaatgtgtgcctgttttgcctcga	1605
QY	1583	aaatttaataaactagaaagatagaaagagggaatagatccttccaatatcaagaag	1642
Db	1606	aagcttgcacaaacagaaagaaagaaagagattgattatatacaaaagccctgaagt	1665
QY	1643	acctga	1648
Db	1666	gaatga	1671
<p>RESULT 4 AAZ95199 ID AAZ95199 standard; DNA; 2092 BP.</p>			
AC	AAZ95199;		
XX	05-JUN-2000 (first entry)		
DE	Human UDP-glucuronosyltransferase 2B4 nucleotide sequence.		
XX	UDP-glucuronosyltransferase 2B4; UGT2B4; polymorphism; metabolism; SNPs.		
KW	drug interaction; detect; human; single nucleotide polymorphism; ds.		

XX	Homo sapiens.
OS	
PN	WO200006776-A1.
XX	
PD	10-FEB-2000.
XX	
PF	22-JUL-1999; 99MO-US16675.
XX	
PR	28-JUL-1998; 98US-0094391.
XX	
PA	(AXYS-) AXYS PHARM INC.
XX	
PI	Galvin M, Miller A, Penny L, Riedy M;
DR	WPI: 2000-195321/17.
P	P-Psdb; AAU78933.
PT	
PP	Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
PT	genotyping individuals to predict rate of metabolism of substrates and
PT	for identifying potential drug interactions
XX	
PS	Disclosure: Page 34-36; 72pp; English.
CC	
CC	This sequence represents the human UDP-glucuronosyltransferase 2B4
CC	(UGT2B4) gene. UDP-glucuronosyltransferase (UGTs) are a family of
CC	enzymes that catalyse the glucuronic acid conjugation of a wide range of
CC	endogenous and exogenous substrates. The UGT2B gene subfamily encode
CC	steroid metabolizing isoforms in the liver. Alteration of the expression
CC	or function of UGTs may effect drug metabolism. The invention relates to
CC	non-chromosomal nucleic acid molecules, which comprise human UGT2B
CC	sequence polymorphisms (see AAZ95051-295110). Probes which detect the
CC	UGT2B locus polymorphisms can be used to detect altered UGT2B metabolism
CC	of a substrate in an individual. The nucleic acid molecules comprising a
CC	human UGT2B sequence polymorphism can be used in screening assays for
CC	genotyping individuals, also to predict their rate of metabolism of
CC	UGT2B substrate, potential drug-drug interactions and adverse side
CC	effects. The polymorphisms can be used as single nucleotide polymorphisms
CC	(SNPs) for detecting genetic linkage related to phenotypic variation in
CC	activity or expression of UGT2B protein. The polymorphism containing
CC	nucleic acid molecules may also be used for generating genetically
CC	modified non-human animals and for obtaining site specific gene
CC	modification in cell lines.
SO	
	Sequence 2092 BP: 639 A; 398 C; 438 G; 617 T; 0 other;
	Query Match 27.8%; Score 766; DB 21; Length 2092;
	Best Local Similarity 68.7%; Pred. No. 2,3e-146;
	Matches 1102; Conservative 0; Mismatches 490; Indels 12; Gaps 3;
OY	34 catcatgaggtctgaacagtcaagcttcgtgatatttgcctccgcagccttttgt---gt 90
DB	
OY	31 catcgagatgctatagaatggacttcagctctcttcttgctgatacagcctgaagctgtaact 90
OY	91 tggcgtgtgattctctggaaagctcctggtgcccctgtgacatgagcattgcttaa 150
DB	
OY	91 tagctctgggatttcttgaaaaggctcgtgtggtggcccaagaatcaccgacttgatgaa 150
OY	151 tgtcaagtcattctagaagagctcatatgtgagagccatgtagtaaacagtatnacta 210
DB	
OY	151 tataaagaacaatccctgtagaacttgcacagagagctcatgagtgacctatgtgcac 210
OY	211 ctcaaagccttgtaattgactlaaaggaagcctctgcatgtaaatgtaggtgtcca 270
DB	
OY	211 ttcagcttcattcttcttcgaccacaagccaccactactctaatttgaagttatcc 270
OY	271 tatgccacagcagacagaagaataatgtaatttgttgcacctagctctca-----a 324
DB	
OY	271 tgtaactttaactaaactagatttgagagatatattcaagcagcgtgttaagaagtggc 330
OY	325 tgccttgccaggtcatcaaacctggcacaatcagttataaataatgattttttgttga 384

Db 331 agaactccaaaagacacatlltggatcatatttccaaagtaacaagaatcatgtgac 390  
 Qy 385 aataagaggaacttaaaaaatgatgtgagagcttcttctcaactcagacgttatga 444  
 Db 391 attaatgacactactagaagaattctctgtataagatatagttcaataagaactatga 450  
 Qy 445 gaagctacagaagaacacactacgtatgtatcttaagacctgtgattccctgtgaga 504  
 Db 451 gaaactacagagatcagaattgtatgttctcttcagatgtgttcccttctgtgaga 510  
 Qy 505 cctgtgtcgtatgt 564  
 Db 511 gctgtgtcgtcgtatgtatgaataacacacttctgtcagacctcgtctcgtgcta 570  
 Qy 565 caatatggagcgaagctgttggaaactccagctccacttccctcgtatgtacgttccat 624  
 Db 571 cgcaattgaaaagcagtagtgagagacttctgtccctcctcctcgtatgtgttat 630  
 Qy 625 gacaggaactaacagacgaatgaccttctgtgaaagagtaaaaaatcattgcttccag 684  
 Db 631 gtcagaactaagtgaccaaagacttctcaatagagaggttaaaaaatcattgcttccag 690  
 Qy 685 ttgttccacttctgtatcagactatcattcatttgggaaagatttatagtatga 744  
 Db 691 ttatttgaatttgttccaaatatttgatagaagaagtgtgatacgttctcagtgta 750  
 Qy 745 ggcatttggagagccacacattatgtgagactgtggaagaactgagatagctat 804  
 Db 751 agtcttggagagacccactacgtatctgtgagacaaatgtacataatgtgctat 810  
 Qy 805 acgaacatattggagatttggatttctcctcaacacacacacttacttgaatttctg 864  
 Db 811 tcgaactactcgtgatttctcctcctcctcctcctcctcctcctcctcctcctcct 870  
 Qy 865 aggttgcactgttaaacctcgcgaagcttgcctgaagaatggaatatttctgcagag 924  
 Db 871 aggtctcactgtcaaacctcgcgaagcttgcctgaagaatggaatatttctgcagag 930  
 Qy 925 ttcaaggagagatgtgtatttctgtgttctcctcgtggtcactgttcaaatgttacaga 984  
 Db 931 ctctggagagaaatgt 990  
 Qy 985 agaaagagctaatatcatcttccacgaccttgcgaagctccacagaaggtgtatagag 1044  
 Db 991 agaaagagctaatatcatcttccacgaccttgcgaagctccacagaaggtgtatagag 1050  
 Qy 1045 gtcaaaagagaaac 1104  
 Db 1051 attgatgtggaataaac 1110  
 Qy 1105 ccgaatgatcttctgtgtatcccaaacgaagcttattatcactcagtgatgaatga 1164  
 Db 1111 ccgaatgatcttctgtgtatcccaaacgaagcttattatcactcagtgatgaatga 1170  
 Qy 1165 tggagatcatgaagctatttaccatgtgggtccctatgtgtgtgtgtgtgtgtgtgt 1224  
 Db 1171 tggacatcatgagcaacttaccatgtgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1230  
 Qy 1225 taagcttgaatcatgtccacatgaagcgaagcagcgtgtgaataataactca 1284  
 Db 1231 tcaaacctgataacatgtcacacatgagcgaagcagcgtgtgaattgtgacttcca 1290  
 Qy 1285 aactatgaacaaagcgaagcttactgagggcttggagaaagcattcaccatcctctta 1344  
 Db 1291 caaatgtcgtgtatcagacttactcactgtcagcgaagaagaatgaatgacttcttata 1350  
 Qy 1345 taaagaagaatgtcatgtatgaatcaagaatcacaatgatacactgttaagcccttga 1404  
 Db 1351 taaaagaatgtcatgtatgaatcaagaatcacaatgatacacaagcgaagcccttga 1410  
 Qy 1405 tcgagcagctctctgtgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1464  
 Db 1411 tcgagcagctctctgtgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1470

Qy 1465 agcttccatbaccctacactgttcccaagactactatagatgtatgtgtgtgtgtgt 1524  
 Db 1471 tgaagccagcagcactcactcactgttcccaagactactatagatgtatgtgtgtgtgtgt 1530  
 Qy 1525 gaccctgt 1584  
 Db 1531 gaccctgt 1587  
 Qy 1585 attaatgaacatagaagaatagaagaaggaatagaatcttc 1628  
 Db 1588 gttcgttgaacaggaagaaggaagaagaatagaatcgtc 1631  
 RESULT 5  
 AAZ95200  
 ID AAZ95200 standard; DNA: 1854 BP.  
 AC AAZ95200;  
 XX  
 DT 05-JUN-2000 (first entry)  
 DE Human UDP-glucuronosyltransferase 2B7 nucleotide sequence.  
 XX  
 KW UDP-glucuronosyltransferase 2B7; UGT2B7; polymorphism; metabolism; SNPs;  
 KM drug interaction; detect; human; single nucleotide polymorphism; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200006776-A1.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PE 22-JUL-1999; 99WO-US16675.  
 XX  
 PR 28-JUL-1998; 98US-0094391.  
 XX  
 PA (AXYS-) AXYS PHARM INC.  
 PI Galvin M, Miller A, Penny L, Riedy M;  
 XX  
 DR MPI: 2000-195321/17.  
 DR P-PSDB; AAT78934.  
 PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for  
 PT genotyping individuals to predict rate of metabolism of substrates and  
 PT for identifying potential drug interactions  
 XX  
 PS Disclosure: Page 41-44; 72pp; English.  
 CC This sequence represents the human UDP-glucuronosyltransferase 2B7  
 CC (UGT2B7) gene. UDP-glucuronosyltransferase (UGTs) are a family of  
 CC enzymes that catalyze the glucuronic acid conjugation of a wide range of  
 CC endogenous and exogenous substrates. The UGT2B gene subfamily encode  
 CC steroid metabolizing isoforms in the liver. Alteration of the expression  
 CC or function of UGTs may effect drug metabolism. The invention relates to  
 CC non-chromosomal nucleic acid molecules, which comprise human UGT2B  
 CC sequence polymorphisms (see AAZ95051-295110). Probes which detect the  
 CC UGT2B locus polymorphisms can be used to detect altered UGT2B metabolism  
 CC of a substrate in an individual. The nucleic acid molecules comprising a  
 CC human UGT2B sequence polymorphism can be used in screening assays for  
 CC genotyping individuals, also to predict their rate of metabolism of  
 CC UGT2B substrate, potential drug-drug interactions and adverse side  
 CC effects. The polymorphisms can be used as single nucleotide polymorphisms  
 CC (SNPs) for detecting genetic linkage related to phenotypic variation in  
 CC activity or expression of UGT2B protein. The polymorphism containing  
 CC nucleic acid molecules may also be used for generating genetically  
 CC modified non-human animals and for obtaining site specific gene  
 CC modification in cell lines.  
 XX  
 XX Sequence 1854 BP; 572 A; 338 C; 392 G; 552 T; 0 other;

Query Match 27.2%; Score 750; DB 21; Length 1854;  
 Best Local Similarity 67.68; Pred. NO. 3.9e-143;  
 Matches 1085; Conservative 0; Mismatches 510; Indels 9; Gaps 2;

34 catatgaagctcagacagcagcttgccttgcctcctcgcagctctcctcgtgtg  
 8 caccaggaatgctcgtgaagatgagcttcagtaatttgcctataaactgagcttgcctt  
 91 tggctgtgactcgtgaggaagctcgtgtgctgagcctgtgaacatgagccattgcttaa  
 68 taagctcgtggaatgtggaagagctgctgtgagcagcagaataacacatgagctgaa  
 151 tgcgaagctatctcctaaagctcctacatgagagccatgaggttaacatgactca  
 128 tataaagacatccctggaatgagcttaccagaggtcagatgaggtgactgactgcatc  
 211 ctcaaacctcgttactacaggaagcctcctgccttgccttgaatttgaaggtccca  
 188 ttacgctcctatcttcttgcctcccaacatccgcctccttaaatgtaaattatcc  
 271 tatgcacagagacagacagagaanaatgaataattgttgcctgactgctga-----a  
 248 cacacttactaaactgattgaggaatttcacatgacatgcaacagattgaagatgctc  
 325 tgccttgcagcttaccacccgtgcaatcagttataaataatgattttttgttga  
 308 agaccttccaaagatcatttgccttataatttccaaagtaaggaataatcagttcaat  
 385 aataaagagactttaaataatgattgtgagagcttatacacaagcagctatgaa  
 368 attgtgacataactagaaagttcgttaagaatgatttcaataaagaataattatgaa  
 445 gaagctacaggaacacacactacatgatacgtatagaccctgtgactccctgtgaga  
 428 aaaaagtaagagcaagatttgcagctatttgcagatgttatttccctgtagtga  
 505 ccgtgagctgagattgtgtcagctccttgccttgcctacacttgaatttcgtagaag  
 488 gctgtgctgagcttacttaacatacccttgcctgtgatacgtcctcctcctgtgcta  
 565 caatatgagcgaagcgtgtggaactcagcctcctcctcctcctcctcctcctcctcct  
 548 cacttttgaagaagcatagtgagatttatttccctcctcctcctcctcctcctcctcct  
 625 gacagagacacagacagacatgaccttctggaagaagtaaaaaatcaatgcttcaat  
 608 gtcaaatatcagatcaatgacattcattcagagaggttaaaaaatgatactgctc  
 685 ttgttccacttctggttcaagatacagctatcatttgcctggaagaagtttataatga  
 668 ttacttgcacttgccttgcgaataattgacatgaaagatgagatcagtttataatgaa  
 745 ggcactggaagccacacatattgtgagactgtggaagaagcagatagatgacat  
 728 agttctaggaagaccacacatattctgagacatgtggaagaagcagatagatgacat  
 805 acgaacatattggaatttgaatttctcctcaacacataacacactaacttgaattgtg  
 788 tgcgaactcctggaatttgcatttgcataccacttccacacttgcgaattgttgcctg  
 865 aggaatgactgtaaacctcgaagcttgcctgaagaagatgaaatgttgcagag  
 848 aggactcactgcaaacctcgaagcttgcctgaagaagatgaaatgttgcagag  
 925 ttcaaggaagatgtattgtgttctcctggttgcctggttcaactgttcaaaatgtaaga  
 908 ctctggaagaaatgtgtgtgtgttctcctggttgcctggttcaactgttcaaaatgtaaga  
 985 agaaagagcctaatactatcgtcctcagccctgtccagatccacagagaagtgatagag  
 968 agaaagagcctaatactatcgtcctcagccctgtccagatccacagagaagtgatagag  
 1045 gtcaaaagaaaaaacatccacatagagacataatccctgctgtatgattgataacc

1028 atttgaaggaaataaacacacataccttaggttccataactcgcgtctcacaagtgatacc  
 1105 ccgaatgatcttcttgcctatcccaaaccaagccttatacctcagatgttgaatgaa  
 1088 ccgaatgactcttgcctatcccaaaccaagccttatacctcagatgttgaatgaa  
 1165 tggatctatgaagcttatacctcagatgttgcctcctcagatgttgaatgaa  
 1148 tggatctatgaagcctatcccaaaccaagccttatacctcagatgttgaatgaa  
 1225 tcaagctgataacatgctcagatgttgcctcctcagatgttgaatgaa  
 1208 tcaagctgataacatgctcagatgttgcctcctcagatgttgaatgaa  
 1285 aactatgacagcgaagatttactgaggttgcctcctcagatgttgaatgaa  
 1268 cacaatgtcagatcagacacttgcctcagatgttgcctcctcagatgttgaatgaa  
 1345 taaagaaatgctatgagattatcaagaatccacatgatacctgttaagccctaga  
 1328 taaagaaatgctatgagattatcaagaatccacatgatacctgttaagccctaga  
 1405 tcaagcagcttctcagatgttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1388 tcaagcagcttctcagatgttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1465 agctgacacagcctcagatgttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1448 tgaagcagcagcactcagatgttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1525 gacctgtgtggaagcctatatttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1508 gctcgtgtggaagcctatatttgcctcctcagatgttgcctcctcagatgttgaatgaa  
 1585 attataaagcagatgagaaagaggaaggaaggaatgatacttgcctcctcagatgttgaatgaa  
 1568 gttgtcagaaagcagaaaggaaggaaggaatgatacttgcctcctcagatgttgaatgaa

RESULT 6  
 AAV87412  
 ID AAV87412 standard; cDNA; 515 BP.  
 XX  
 AC AAV87412;  
 XX  
 DT 27-APR-1999 (first entry)  
 XX  
 DE EST clone BR77.  
 XX  
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 OS Homo sapiens.  
 XX  
 PN J09845435-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98MO-US06954.  
 XX  
 PR 10-APR-1997; 97US-0835913.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 XX  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
 PI Racie LA, Spaulding V, Treacy M;  
 XX  
 DR WPI; 1999-070076/06.  
 XX  
 PT New polynucleotides encoding human secreted proteins - derived from



PR e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PR ovary, pituitary, retina and colon cDNA libraries  
 XX  
 PS Claim 1; Page 556; 633pp; English.

Claim 1; Page 556; 633pp; English.

This sequence represents an expressed sequence tag (EST), and is a polynucleotide of the invention. The polynucleotides of the invention all secreted EST sequences isolated from a variety of human tissue sources. The EST sequences and proteins encoded by them are predicted to have useful biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, immune stimulating or suppressing activity, haematopoietic regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, tumour inhibition activity. The EST sequences are also stated to be useful for gene therapy.

SQ Sequence 515 BP; 148 A; 98 C; 122 G; 147 T; 0 other;

Query Match	17.68;	Score	486.4;	DB	20;	Length	515;
Best Local Similarity	99.8%;	Pred. No.	8e-90;				
Matches	487;	Conservative	0;	Mismatches	1;	Indels	0;
						Gaps	0;

OY	27	gaatcgccatcaatgagctcgtacaaagtcagcttgtagtattctcgtccctgacgtctct	86
Db	22	gaacttgcacatcatgaggtcgttcgaagctcagcttctgtagtattctcgtccctgacgtctct	81
OY	87	ggttggtcgttgatctctgttggaaagccctgggttgcccttggaacatgaagccattggc	146
Db	82	gggttggtcgttgatctctgttggaaagccctgggttgcccttggaacatgaagccattggc	141
OY	147	ttaatgtcaaggtcaatcttctaagaagctcctacagtgtgaagggccatgaagtaacagtatga	206
Db	142	ttaatgtcaaggtcaatcttctaagaaggtccacatagtgagagggccatgaagtaacagtatga	201
OY	207	ctccatccaagccttcgtttaattgacctacagaagcccttcgcattgaatttgaagtgg	266
Db	202	ctccatccaagccttcgtttaattgacctacagaagcccttcgcattgaatttgaagtgg	261
OY	267	tccataatgccacaggaacagaacagaagaataatgttctgacactagctctgatatg	326
Db	262	tccataatgccacaggaacagaacagaagaataatgttctgacactagctctgatatg	321
OY	327	tcttgcacaggtcttaccacccctgggaatcgaattataaataatgaatttttttgattgaa	386
Db	322	tcttgcacaggtcttaccacccctgggaatcgaattataaataatgaatttttttgattgaa	381
OY	387	taagaggaactttaaaaaatgtatgtgtgtagaggttatactaaatcaagcgttatgaaga	446
Db	382	taagaggaactttaaaaaatgtatgtgtgtagaggttatactaaatcaagcgttatgaaga	441
OY	447	agctaacaggaacccaactacagatgaattgtttaagaagccctgtgatccctgtggagacc	506
Db	442	agctaacaggaacccaactacagatgaattgtttaagaagccctgtgatccctgtggagacc	501
OY	507	tgatgtgct 514	
Db	502	tgatagct 509	

## RESULT 7

ID AAC03286 standard; cDNA; 350 BP.

AC AAC03286;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3284.

XX	
KW	Human; 5' EST; expressed sequence tag; secreted protein; cdna isolation,
KW	gene therapy; chromosome mapping; ss.
XX	

OS Homo sapiens

PN EP1033401-A2

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.  
VV

PA (GEST ) GENSET.  
yy

PI Dumas Milne Edwards J, Duclert A, Giordano J;  
xy

DR WPI; 2000-500381/45.  
DR P-BDDP; 2000-500380.

XX  
DT

Claim 1: SEQ ID 3284, 71pp + CD-ROM, English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or poly(A) RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of cDNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

SD Sequence 350 BP; 108 A; 69 C; 77 G; 96 T; 0 other;

Query Match	12.7%	Score 350;	DB 21;	Length 350;
Best Local Similarity	100.0%	Pred. No. 3.3e-62;		
Matches 350; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	791	gattataggttaacgaacatttggatttgaatttcctccaacataccaactaac	850
Db	1	gagatacaggtcaatacgaatacatatvggatttggatttcctccaacataccaactaac	60
QY	851	tttgaatttcttgagagattgcactgtaaacctgcgaagcttttgctaaggaataatgaa	910
Db	61	tttgaatttcttgagagattgcactgtaaacctgcgaagcttttgctaaggaataatgaa	120
QY	911	aatttgcctcagaagttccagggagaagtctgatttggttttctctgggtgcactgttt	970
Db	121	aatttgcctcagaagttccagggagaagtctgatttggttttctctgggtgcactgttt	180
QY	971	caaaatgttaacagaagaaaaaggtcaatatcatctgttccagcccttgcgcagaatccccaag	1030
Db	181	caaaatgttaacagaagaaaaaggtcaatatcatctgttccagcccttgcgcagaatccccaag	240
QY	1031	aaaggttttatgggggtacaaagaagaaaaaacatccatgaaggagccaatattctgtgtg	1090
Db	241	aaaggttttatgggggtacaaagaagaaaaaacatccatgaaggagccaatattctgtgtg	300
QY	1091	tatgatattgataccccaagaatgatccttcttgtgtcatccccaacccaagaagc	1140
Db	301	tatgatattgataccccaagaatgatccttcttgtgtcatccccaacccaagaagc	350



```
RESULT 8
AAQ27370 ID AAQ27370 standard; cDNA, 2368 BP.
XX AC AAQ27370;
XX 27-JAN-1993 (first entry)
XX DE HUG-Br2.
XX Biliirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
XX monoglucuronide; diglucuronide; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 30..801
XX FT FT /*tag= a
XX FT FT 2347..2352
XX FT FT /*tag= b
XX FT FT /*number= 1
XX FT FT 2355..2360
XX FT FT /*tag= c
XX FT FT /number= 2
XX PN WO9212987-A.
XX PD 06-AUG-1992.
XX PF 10-JAN-1992; 92MO-US00282.
XX PR 10-JAN-1991; 91US-0639453.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX PI Owens IS, Ritter JK.
XX DR WPI: 1992-284593/34.
XX DR P-PSDB; AAR26154.
XX PT Isolated gene locus UGT1, DNA segments and diagnostic probes -
XX PT for diagnosing Gilbert's disease and Crigler-Najjar syndrome
XX PT types I and II
XX PS Disclosure: Fig 9A-I; 99pp; English.
XX CC Two human liver bilirubin UDP-glucuronosyltransferase cDNAs have
XX CC been isolated. They are referred to as HUGBr1 (AAQ27369) and HUGBr2
XX CC (AAQ27370) (Ritter, et al., J. Biol. Chem. 266:1043-1047 (1991)) and,
XX CC upon expression individually in COS-1 cells, encode isoforms that
XX CC catalyze the formation of the two bilirubin monoglucuronides and
XX CC the diglucuronide.
XX CC The cDNAs contain identical 3' ends (1469 bp in length) to each
XX CC other and to that of the human phenol transferase cDNA, HUGP1
XX CC (Harding et al., Proc. Natl. Aca. Sci. USA 85:8281 (1988)).
XX CC In contrast, they have unique 5' ends.
XX SQ Sequence 2368 BP; 609 A; 532 C; 566 G; 661 T; 0 other;

Query Match 11.5%; Score 317.8; DB 13; Length 2368;
Best Local Similarity 52.0%; Pred. No. 1.7e-55;
Matches 790; Conservative 0; Mismatches 717; Indels 12; Gaps 3;
```

```
QY 188 catgagtaacagatttgactcaatcaagaagccttgcttaattgactacagaagcctct 247
DB 195 caccaggcggtggtctccaccaccaagaagtgaaatgacatcaagaagaatttttc 254
QY 248 gcatgaaatttgaggtggtccatagtcacagagacagacaagaagaattgaatatt 307
DB 255 accctgacagcctatgctgttccatgaccag-----aaggaatttgatcggttaag 308
QY 308 gttagactagctcgaatgcttgcaggcttatacaactgcgaatcagttataaatta 367
DB 309 ctgggtacacccaagggtcttctgaacagaacatctctctgaagagatattctagaagt 368
QY 368 aatgatttttcttgtaataagagaacttcaaaatgatgtgtagagcttatctac 427
DB 369 atggcaattatgaacaatgatacttgccttcaataggtgtgtgtagactgcat 428
QY 428 aatcagacgcttatgaagaagctacagagaacaactacagatgtaattgattagacct 487
DB 429 aatgagcgccctgatacagagacactgatactctcttgatggtgtttaacagagccc 488
QY 488 gtgattcccttgtagaagcctgatagtgctgagtgttgcaagtcaccttgtgtacact 547
DB 489 gttaaacctgcgggcggtgtgctgtaagtaacctgtcattcctgctgtgtt---ttt 545
QY 548 agaattctgtaggaggaatatagtagcggaagctgtggaagaacttccagctccacttc 607
DB 546 tggaggtacatccatccatgtaacttaagggacacagatgltccaaatccctctcc 605
QY 608 tatgtacctgtcctatgacagagactaacaagaatgaccttctctggaagaagtaaa 667
DB 606 tatattccctaagtactaagcagacaaatlcagacacaatgacatctctggaagaag 665
QY 668 aattcaatgcttcaagtttcttccactctgagatcagatgaactaactaatttttg 727
DB 666 aac--atgcttaacctctgacctgtcccttcaacttgaactttctgaccttat 722
QY 728 gaagagtttatagtaagcattaggaagccacatacttatgtagagcttgggaaga 787
DB 723 gcaagctctgctctgagcttttccagagaaggtgtaactgcgtccagaactattgaact 782
QY 788 gctgagatagctgaatacagaacatacttggaatttggaatttccctccacataccacct 847
DB 783 gactcgtctgctgttgtaagaagtacttgtaagaattacacttaggacatcatgccc 842
QY 848 aacttgaagttgttggaagatgtaactgaaacctgcacaagcttgccttaaggaatg 907
DB 843 aatatgtcttcaatgtgggcatcaactgttccacagggagagccactatcctagaattt 902
QY 908 gaaaatttgcacaggttccagggaagatgtgattgtgttctctctgaggaactg 967
DB 903 gaagcctacatlaatgcttcttggaagacatggaattgtgttctctcttgaatcaatg 962
QY 968 ttcaaaaatgctacagaagaagaagctaaatcatcatgcttccagcccttgcacaccca 1027
DB 963 gctcgaagaattccagagaagaagatgtaagcttgatgcttgggcaaaatccct 1022
QY 1028 cagaaggttatgtgaggtacagaagaacaaacatccacattaggaagcaatactcg 1087
DB 1023 cagacagctcgtgtggtgactgactggaacccagcaatccttgcgacacacagata 1082
QY 1088 ctgtatgattgataccccaagatgattcttctgtgacacccaacaaacatttatc 1147
DB 1083 ctgttaagtggtctacccaagaacgactgccttgcgtgcacccgagtgccttatcc 1142
QY 1148 atcatatgtagaataatgagatcataagaactatttaccatgggggttccatgtggtga 1207
DB 1143 acccatgctggttcccatcagtggtttatgaagacatatagtaagtggttccatggtatg 1202
QY 1208 gtcccatatttgtagatcagctgataacatagctcacatgaagggcaagagcagct 1267
DB 1203 atgcctcttggtgtatcagatgtaagcgaatgtaagcgaatgaagagagctgta 1262
QY 1268 gtagaataaacttcaaaaactatgacaagcgaaagatttactgaaggctttagaacaagtc 1327
```

```
Db 1263 gtgacctgagatgtctggaatgactctcgaatttagaataatgtcctcaaaagcagtc 1322
Oy 1328 attaccgattccctcttaataagaaatgataatgattatcaagaattccacatgacaa 1387
Db 1323 atcaatgacaaaattacaaagaaacatcatgcgctccacgctccacaaagaccgc 1382
Oy 1388 cctgtaagcccccatagcagatctctgtatgcagatttctcagcgcccaagaagga 1447
Db 1383 ccggtgagcgcctgagaccctgcccgtctctgtgtgaggttctgagtgagcgcaagagc 1442
Oy 1448 gccaaagcctgagatcagctgcccacatgacctcagctgtctcagcactactatagat 1507
Db 1443 ggcgcacacccctgcccgcagccacagaccctcagctgacagatccatctctggac 1502
Oy 1508 gtgattggtctcctgcctgactctgtggcaactgctataatctgttcacaaaatgttt 1567
Db 1503 gtgattggtctcctctgcccgtgctgagcagtgccctcatcaacttaaatgtgt 1562
Oy 1568 ttattctctgtcaaaaat 1586
Db 1563 gcttatgctacccgaaat 1581
```

## RESULT 9

AA027369  
ID AA027369 standard; cDNA; 2351 BP.

AC AA027369;

DT 27-JAN-1993 (first entry)

XX HUG-Brl.

KW Billirubin; UDP-glucuronosyltransferase; HUGBrl; HUGBr2;  
monoglucuronide; diglucuronide; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 16..784

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

FT polyA\_signal

```
DB 924 gtttctcttggatcaatggtctcagaattccagagaagaagctatgacattgc 983
OY 1006 ttccgaccttgcagatcccaagaagtggtatgaggtacaaagaaaaaacatc 1065
DB 984 tgaacttgggcaaaatccctcagacagctcgtggtgacactggaacccgacatc 1043
OY 1066 caacttggagccaatcgcgtgtatgtttgtagtccccagatattctctgttca 1125
DB 1044 gaacttgcgaacaacaacagatactgtttaagttggtaccccaaacgactcgtgttca 1103
OY 1126 tcccaaaccaaaagctttatcactcagtggtgatatgagtgagtcataagatatta 1185
DB 1104 ccgagtgcacccggtccttatccaccatgctgttcccatggtgtttatagaacatata 1163
OY 1186 ccattggttcccatggttggagttcccatatttggtagtcagcttataacatagctca 1245
DB 1164 caattggttcccatggttggatgaccccttgttggtagtcagatggaacatgcaacgcg 1223
OY 1246 catgaagcccaaaaggaacagctgtagaataaacttcaaaactatggcaagcgaagatt 1305
DB 1224 catggaactcaaaaggagctggaagccttgaaactgttcttgaatccttcgaagattc 1283
OY 1306 actgagggcttggagaacagctatcaccgattcctctataaagaatgctatgagatt 1365
DB 1284 agaaaatgctcaaaaagcagctatcacaatgacaagaagtaacatcatcgtgcct 1343
OY 1366 atcaagaattcacaatgataacctgttaagccccctagatcgagagcttcttgatcga 1425
DB 1344 ctccagcttccacaagaagccgcgtgtgagccgcttgactggtcgttcttctgtgga 1403
OY 1426 gtttgcatacgcccaaaagagcaagcagccttgcagtcacagctccatgacccactctg 1485
DB 1404 gttgtgtatgaagacaaaggcgcgcgcacacctgcgcgcgcgcgcgcgcgcgcgcgcgc 1463
OY 1486 gtcccaagcactactataatgattgattggttcttcttgcagccgtgttggcaactgtcat 1545
DB 1464 gtaccagctacatctccttgcagctgattgttcttcttgcgcgtgttgcgaagcgcc 1523
OY 1546 attcctgttcaaaaatggtttttatttcttcttcttcaaaaat 1586
DB 1524 ctccatcactttaaatgttgtgttattgtgtctacccgaat 1564

RESULT 10
AAZ45118
ID AAZ45118 standard; DNA; 735 BP.
XX
AC AAZ45118;
XX
DT 28-FEB-2000 (first entry)
XX
DE UDP-glucuronosyltransferase 1 (UGT1) exons 2-5 nucleotide sequence.
XX
KW uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
KW unconjugated hyperbilirubinemia; drug metabolism; transgenic animal;
KW pharmacogenetic screening; diagnosis; ss.
XX
OS Homo sapiens.
XX
PN MO957322-A2.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1999; 99MO-US09702.
XX
PR 07-MAY-1998; 98US-0084807.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Penny L, Galvin M;
XX
DR WPI; 2000-052981/04.
```

```
DR P-PSDB; AAY57100.
XX
XX New nucleic acid representing polymorphisms in the human uridine
PT diphosphate glucuronosyltransferase gene, used for diagnosis and
PT evaluation of drug metabolism
XX
XX Examples: Page 43-44; 63pp; English.
XX
XX Nucleotide sequences AAZ45110-45118 are exons 1A-1J of the human
CC uridine diphosphate-glucuronosyltransferase 1 (UGT1) gene. The UGTs are a
CC family of enzymes that catalyze the glucuronic acid conjugation of a wide
CC range of endogenous and exogenous substrates including phenols,
CC alcohols, amines and fatty acids. Many of the reactions catalysed by UGTs
CC result in toxic substances being converted to compounds which are more
CC water soluble and are excreted. The invention relates to and identifies
CC UGT1 polymorphisms (AAZ45004-245041). The polymorphism sequences are
CC useful as probes for detecting UGT1 locus polymorphisms, indicative of
CC altered UGT1 expression or activity. These polymorphisms are associated
CC with Crigler-Najjar and Gilbert syndromes (unconjugated
CC hyperbilirubinemia) and drug metabolism. The genotyping of the UGT1 gene
CC is used to predict the rate of metabolism of UGT1 substrates, possible
CC drug-drug interactions and adverse side effects (i.e. to optimize drug
CC dosage), and to screen for diseases caused by exposure to toxins and to
CC study the effects of polymorphisms on enzymatic activity. The UGT1
CC sequences, including polymorphisms, can also be used to produce the
CC corresponding protein (or its fragments) or to generate transgenic
CC animals or modified cells e.g. for pharmacogenetic screening.
XX
XX Sequence 735 BP; 189 A; 179 C; 187 G; 180 T; 0 other.

Query Match 9.9%; Score 272.2; DB 21; Length 735;
Best Local Similarity 62.3%; Pred. No. 2.3e-46;
Matches 427; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

OY 902 gaatggaanaatttgcacagatccagggagaagatggtatggtgttcttctcgtgg 961
DB 1 gaattgaagctcatcaltatgcttcttgcgagaacatggaattggtttcttcttggga 60
OY 962 tcaatggttcaaaatgtctacagaagaagaagctaatatcatcttgcagcccttgcag 1021
DB 61 tcaatggttctcagaatctccagagaagaagctatgtaactgttgcattgttggcaca 120
OY 1022 atcccaagaagtggtatgaggtacaaaggaaaaaaacatccatcattggagccat 1081
DB 121 atccctcagacagctcgttgcgtgtacacttgcacccgacatcgaatcttgcgaacac 180
OY 1082 actgcgtgtatgttgcataccccaagatgatcttcttgcatacccaaaccaagct 1141
DB 181 acgatacttctaagttggtaccccaaaacgactgcgttgcacccgatacgcgtgcc 240
OY 1142 ttatcactcagtggtgaatggtatggtatcataagactatccatgggttccctatg 1201
DB 241 ttatcaccactgtgttcccatggtgtttatgaagaacatagatgaatgggttccatg 300
OY 1202 gtggaggttcccatatttggtagcagcttgaataacatagctacatagaagccaaga 1261
DB 301 gtgatgagcccttgcgttgcagatgagcaaatgacaagcgaatgagactaagga 360
OY 1262 gcaagctgtagaataaacttcaaaactatgcaagcgaaggtttactagaagggtttga 1321
DB 361 gctggaagtgaccctgaatgttcttgaatagctcttcgaagatttgaagaatgctctaa 420
OY 1322 acaatcattacagcttctctataaagaagatgctatgaatatacaagaatccacat 1381
DB 421 gcaatcattacagcttctctataaagaagatgctatgaatatacaagaatccacat 480
OY 1382 gataacctgtaagcccttagatcgagcagcttcttgcagagttgttcatgcgcac 1441
DB 481 gaccgcccgtgtagcgcgttgcagcttgcgttcttgcgtgtgaggttgcgtatgagcgc 540
OY 1442 aaaggagcaagcagcctggtcagctgcgcattacactcacttgcacacactactct 1501
```

Db 541 aaggcgccgacacactgccccgcaagccacagcctcacttggtaccagttaccctcc 600  
Oy 1502 atagtgatgattgggttccctcctgaactcgtgtggaactgtatatcttgttcaaaa 1561  
Db 601 ttgagcgtgattggttctcctccttgccgctgctgacagtggtccctcaacatttaa 660  
Oy 1562 tgttttatttctcctgcacaaaat 1586  
Db 661 tgttgtcttatgctaccggaat 685

RESULT 11  
AAZ95207  
ID AAZ95207 standard; DNA; 2312 BP.  
XX  
AC AAZ95207;  
XX  
DT 05-JUN-2000 (first entry)  
XX  
DE Human UGT2B15 exon 1 nucleotide sequence.  
XX  
KM UDP-glucuronosyltransferase 2B15; UGT2B15; polymorphism; metabolism;  
KM drug interaction; detect; human; single nucleotide polymorphism;  
KM SNPs; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200006776-A1.  
PD 10-FEB-2000.  
XX  
PF 22-JUL-1999; 99WO-US16675.  
XX  
PR 28-JUL-1998; 98US-0094391.  
XX  
PA (AXYS-) AXYS PHARM INC.  
XX  
PI Galvin M, Miller A, Penny L, Riedy M;  
DR WPI; 2000-195321/17.  
XX  
PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for  
PT genotyping individuals to predict rate of metabolism of substrates and  
PT for identifying potential drug interactions  
XX  
PS Example 3; Page 60-61; 72pp; English.  
XX  
CC This sequence represents the nucleotide sequence of exon 1 of the human  
CC UDP-glucuronosyltransferase 2B15 (UGT2B15) gene.  
CC UDP-glucuronosyltransferase (UGTs) are a family of enzymes that catalyze  
CC the glucuronic acid conjugation of a wide range of endogenous and  
CC exogenous substrates. The UGT2B gene subfamily encode steroid  
CC metabolizing isoforms in the liver. Alteration of the expression or  
CC function of UGTs may effect drug metabolism. The invention relates to  
CC non-chromosomal nucleic acid molecules, which comprise human UGT2B  
CC sequence polymorphisms. Probes which detect the UGT2B locus polymorphisms  
CC can be used to detect altered UGT2B metabolism of a substrate in an  
CC individual. The nucleic acid molecules comprising a human UGT2B sequence  
CC polymorphism can be used in screening assays for genotyping individuals,  
CC also to predict their rate of metabolism of UGT2B substrate, potential  
CC drug-drug interactions and adverse side effects. The polymorphisms can be  
CC used as single nucleotide polymorphisms (SNPs) for detecting genetic  
CC linkage related to phenotypic variation in activity or expression of  
CC UGT2B protein. The polymorphism containing nucleic acid molecules may  
CC also be used for generating genetically modified non-human animals and  
CC for obtaining site specific gene modification in cell lines.  
XX  
SQ Sequence 2312 BP; 765 A; 360 C; 414 G; 773 T; 0 other;

Query Match 7.3%; Score 200.2; DB 21; Length 2312;  
Best Local Similarity 56.6%; Pred. No. 1.2e-31;  
Matches 417; Conservative 0; Mismatches 308; Indels 12; Gaps 2;

Oy 35 atcatgagctcgaacagtcacgtcttggtatctctgcctcgcagctctctgt---gtt 91  
Db 696 accagagtgctcctcctgaatgagcgcagctctctcgcctgatacagctcagttgtaactt 755  
Oy 92 ggcctggaattccttggaagaagctcgtgtgccccctgtgacatgagccattggtta 151  
Db 756 agctctggaagctgtggaagagtgctagtggtgccccagataacagcattggaataat 815  
Oy 152 gtcaagtcattcgaagagctcagtagtgagagcctgaggaacgtatttgactaac 211  
Db 816 atgaaagacaatccctggagagcctgttcaagagggcctcagagctgtgttgaacatc 875  
Oy 212 tcaagccttcgttaattgactacaggaagccttcctgcattgaatttgaggtgtccat 271  
Db 876 tcgctctactctgtgcaatgcagtaaatcatcctgtatgaattgaagttatctct 935  
Oy 272 atgcc-----acggagcagacagagaagaataatgaaatattgttgacctagctc 322  
Db 936 acatccttaactaaaatgatttggaaatctcctctgaaaatctcgtatagatgata 995  
Oy 323 aatgcttgccaggtctatcacctggcaatcagttataaataatgatttttgtt 382  
Db 996 tatgtgtttcaaaaaatacatcttggatcatatttccacattacagaagtgtgttg 1055  
Oy 383 gaaataagaggaactttaaaaatgattgtgagagccttaltctaacatcagacgttatg 442  
Db 1056 gaataattatgactacagtaacaaagctctgttaagaatgagtttgaataagaactatg 1115  
Oy 443 aagaagctacaggaacaaactacatgataatgtttagacccctgagctcctgtgga 502  
Db 1116 atgaaactacaagagtgtaaaagtttgatgcatcttcgcagatagcccttaacctgtg 1175  
Oy 503 gacctgattgctgagttgctgacgtcccttctgttccacacttagaattctgtaga 562  
Db 1176 gagctactgctgacatttaacatcccttctgtacagctcttcgtatctctgtgc 1235  
Oy 563 ggcataatgagcgaagctgtggaactccacgtccactcttcctgtgacccgtgct 622  
Db 1236 taccatttgagaagaatggtggaagattctcctccctccctcctgtaacctgtgt 1295  
Oy 623 atgacaggaactacagacagaatgaccttctcggaaagagtaaaaatgaatgattca 682  
Db 1296 atgcagaatlaagtgtaacatgatttcatgagagggatlaaaaatgatgatacatg 1355  
Oy 683 gtttgcacctctggtatcaggtatgactatcatttttgggaaggtttatagt 742  
Db 1356 cttaatttgaccttggttccaattatgatactgaaagaatggaagcagttttagt 1415  
Oy 743 aagcattagaagagcc 759  
Db 1416 gaagttcagtagtaagtc 1432

RESULT 12  
AAF58252  
ID AAF58252 standard; DNA; 936 BP.  
XX  
AC AAF58252;  
XX  
DT 24-APR-2001 (first entry)  
XX  
DE Oligonucleotide D1835.  
XX  
KM Electron-transfer group; ETM; mismatch; genotyping;  
KM gene expression; ss.  
XX  
OS Synthetic.  
XX  
PN WO200107665-A2.  
PD 01-FEB-2001.

Oy	2433	taatgctcattctttaaaatgatgaataaccataattcttatcttataatcaagta	2492
Db	664	www.....	723
Oy	2493	taattactgtagaanaataaagatgcttgctctgaagttaaaaaaaaaaaaaa	2552
Db	724	www.....	783
Oy	2553	aa 2554	
Db	784	ww 785	
 RESULT_13			
ID	AAF58254		
AC	AAF58254 standard; DNA, 936 BP.		
XX			
AC	AAF58254;		
XX			
DT	24-APR-2001 (first entry)		
XX			
DE	Oligonucleotide D1875.		
XX			
KW	Electron-transfer group; ETM; mismatch; genotyping; gene expression; ss.		
XX	Synthetic.		
OS			
PV	WO200107665-A2.		
XX			
PD	01-FEB-2001.		
XX			
PF	26-JUL-2000; 2000WO-US20476.		
XX			
PR	26-JUL-1999; 99US-0145695. 17-MAR-2000; 2000US-0190259.		
XX			
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
XX			
PI	Umek RM;		
DR	WPI; 2001-159728/16.		
XX			
PT	Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -		
PS			
XX	Example 6; Page 127; 159pp; English.		
CC			
CC	The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.		
CC			
XX			
SQ	Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;		
 Query Match            7.1%; Score 196.4; DB 22; Length 936; Best Local Similarity   0.6%; Pred. No. 5.6e-31; Matches     5; Conservative   548; Mismatches   229; Indels     0; Gaps       0;			
Oy	1773	tatatcatcttcgtgcgaattccaggtagtgcctaactcttcctaattgtg	1832
Db	4	www.....	63
Oy	1833	acaacaagcatgaatacataatttcattctatctgataccaactgtttccatgacgt	1892
Db	64	www.....	123



```
QY 2493 taattactctgtagaataaagagatgctgttctgaaagtaaaaaaaaaaaaaa 2552
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 724 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 783
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2553 aa 2554
      : :
DB 784 ww 785

RESULT 15
ID AAF58259
AF AAF58259 standard; DNA; 936 BP.
AC AAF58259.
DT 24-APR-2001 (first entry)
DE Oligonucleotide D2004.
KW Electron-transfer group; ETM; mismatch; genotyping;
   gene expression; ss.
OS Synthetic.
PN WO200107665-A2.
PD 01-FEB-2001.
PF 26-JUL-2000; 2000MO-US20476.
PR 26-JUL-1999; 99US-0145695.
   17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
PI Umek RM:
DR WPI; 2001-159728/16.
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
PS Example 6; Page 128; 159pp; English.
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 7.1% Score 196.4; DB 22; Length 936;
Best Local Similarity 0.6%; Pred. NO. 5.6e-31;
Matches 5; Conservative 548; Mismatches 229; Indels 0; Gaps 0;

QY 1773 tatatctatcttctgtgatttctcttaagtggtgcttactcttctcacttg 1832
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 4 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1833 acacaagacatgaatacatcattctctatctgatacactgtttccatgact 1892
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1893 cattaactcttaacctaaagtgaagggaacctgcaatagctgactctgtgttg 1952
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1953 cacaaacacatgattgaagaataaataatgtataaattccaaaattcagtaaacac 2012
```

```
DB 184 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 243
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2013 acaaatcaatgaagcattctcatgaacttagctgttaltgtaacataatgatttct 2072
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 244 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2073 ttccaatttaataagcccttccatcacaccagcattactgctcagaataatgac 2132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 304 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2133 taaaatgacgatagggcattacactcagaatagtgtctatatccacatacctc 2192
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 364 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2193 tagatgcatagccctacattctgcacactcaactgaactttttgtgttctgat 2252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 424 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 483
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2253 gataaatagacagttcttattatgtctcctaataataaagaactgaatttcttac 2312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 484 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 543
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2313 atagaagaatgtccataagaatalccaagltaaacagattattlgagataagtaacac 2372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 544 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 603
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2373 tagaataatgtatgttaattctgatttataaatttaattgataagtaacactgat 2432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 604 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 663
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2433 taatgtctattctttaaagtgaatactcataatcttalcctcataccaagta 2492
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 664 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 723
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2493 taattactgtagaataaataaagagatgctgtctgaaagtaaaaaaaaaaaaaa 2552
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 724 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 783
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2553 aa 2554
      : :
DB 784 ww 785
```

Search completed: August 13, 2001, 12:27:18  
Job time: 6449 sec



